Uncertainty Relations for Positive Operator Valued Measures

Serge Massar
Laboratoire d'Information Quantique, C.P. 225, Université Libre de Bruxelles (U.L.B.), Boulevard du Triomphe, 1050 Bruxelles, Belgium

(Dated: March 6, 2007)

How much unavoidable randomness is generated by a Positive Operator Valued Measure (POVM)? We address this question using two complementary approaches. First we study the variance of a variable associated to the POVM outcomes. We illustrate this method by generalizing the inequality proposed by B.-G. Englert for joint measurements of which path / interference visibility in a Mach Zehnder interferometer. Second we study lower bounds on the entropy of the POVM outcomes.

arXiv:quant-ph/0703036v1 5 Mar 2007

Introduction. Richard Feynman famously said that "no one understands quantum mechanics". Nevertheless, practitioners of quantum mechanics have a number of mental pictures that help them reason about the theory. Arguably the most important of them is the uncertainty principle [1], which both gives fundamental insights and provides important quantitative predictions.

The uncertainty principle itself comes in several forms. The best known bounds the variances of two operators $A = \sum_k a_k |a_k\rangle\langle a_k|$ and $B = \sum_l b_l |b_l\rangle\langle b_l|$ measured in the quantum state $|\psi\rangle$. The resulting inequality [2] is:

$$\Delta A \; \Delta B \geq \frac{1}{2} |\langle\psi| \, [A \, , \, B] \, |\psi\rangle| \tag{1}$$

where $\Delta A^2 = \left\langle A^2 \right\rangle - \langle A\rangle$, $\Delta B^2 = \left\langle B^2 \right\rangle - \langle B\rangle^2$. In terms of position and momentum operators it takes the form [3]: $\Delta x \; \Delta p \geq \frac{\hbar}{2}$. An alternative formulation based on entropic uncertainty relations [4] provides bounds on the entropy of the outcomes of measurements of A and B. The version conjectured in [5] and proved in [6] takes the form:

$$H(A) + H(B) \geq -\log_2 \max_{kl} |\langle a_k | b_l\rangle|^2 \tag{2}$$

where $H(A) = -\sum_k p_A(k) \log_2 p_A(k)$ is the Shannon entropy of the probability distribution $p_A(k) = |\langle a_k|\psi\rangle|^2$ and similarly for $H(B)$. In the case of position and momentum measurements the entropic uncertainty relation had been derived earlier in [7]. Entropic uncertainty relations have some conceptual advantages over the relation (1): they are independent of the values a_k, b_l one assigns (often arbitrarily) to the outcomes of the measurements of A and B; and their right hand side is independent of the quantum state $|\psi\rangle$, whereas the right hand side (1) can even vanish although ΔA and ΔB are both positive. Note that in the above formulations the measurements of A and B are mutually exclusive events: they cannot both be carried out.

A conceptually different form of the uncertainty principle concerns measurements that simultaneously estimate two non commuting operators A and B. The precision ΔA and ΔB with which A and B are jointly estimated should obey a constraint similar to eq. (1). Proving such relations for constraints is difficult, see [8, 9, 10, 11, 12, 13, 14, 15, 18, 19] for some of the works in this direction.

Simultaneously estimating two observables cannot be carried out within the usual framework of von Neumann measurements, also called Projector Valued Measures (PVM). Rather it must be formulated within the more general context of Positive Operator Valued Measures (POVM). Formally a POVM $\mathcal{M}$ is described by a set of positive operators that sum to the identity: $\mathcal{M} = \{m_k\}$, $m_k \geq 0$, $\sum_k m_k = \mathbb{1}$.

POVM's play an essential role in measurement theory as they describe measurements affected by noise, fuzzy measurements, measurements that simultaneously estimate two observables. POVM's also play an essential role in quantum information: they are often the measurements which allow the most information to be extracted from a quantum system[20, 21], they are widely used for quantum communication tasks, etc... However POVM's suffer from the fact that (except when some of the POVM elements are projectors) there is no state $|\psi\rangle$ for which the outcome of the POVM is fixed. That is, all most all POVM's contain some inherent uncertainty: their outcomes are affected by unavoidable noise. This raises a fundamental question: why is it that, given their unavoidable noise, POVM's are nevertheless sometimes better than PVM's for information processing tasks? A partial explanation is that the noise of a POVM is more uniformly distributed over Hilbert space: there are many states on which a PVM will give completely random results, whereas many POVM's never give completely random outcomes. But it is unclear whether this is the whole answer. Indeed there at least one very specific context in which the randomness produced by a POVM can be removed: if the same POVM must be carried out on many independent states, then one can devise a collective POVM acting on all the states which is almost a PVM, and which when restricted to a single system, acts as the original POVM[22]. Thus in this very specific context the extra randomness produced by the POVM can be removed.

The inherent randomness of POVM's may also explain some of their limitations. For instance there is to our knowledge no non locality experiment for which POVM's are better than PVM's, and this is probably due to the added noise coming from POVM's. Better understanding the potentialities and limitations of POVM's thus hinges on better understanding the unavoidable noise they add to experiments. In the present work we address this task.

There are many works that address aspects of this question [8, 9, 10, 11, 12, 13, 14, 15, 18, 19], often in the context of joint measurements of position and momentum. But it is clear that a unified approach is necessary that does not focus on the technical difficulties of infinite dimensional spaces, but rather on the conceptual issues involved. The present work aims to fill this gap by following two complementary approaches, similar in spirit to the complementary approaches provided by the Robertson and entropic uncertainty relations mentioned above.

Uncertainty Operator. In our first approach (see also [10, 13]) we suppose that to each POVM element m_k one associates a real value $\mu_k \in \mathbb{R}$. This association of course suffers from the same limitations as the Robertson inequality: the choice of the μ_k is to some extent arbitrary. Different choices will yield different estimates of the uncertainty. Furthermore, it is often natural to associate several values to the same POVM element m_k. We will then have several uncertainties associated with the same POVM.

The expectation value of μ_k is

$$\bar{\mu} = \sum_k \mu_k \langle\psi|m_k|\psi\rangle = \langle\psi|M|\psi\rangle \tag{3}$$

where we introduce the operator

$$M = \sum_k \mu_k m_k \ . \tag{4}$$

The variance of μ_k is

$$\begin{aligned} Var(\mu) &= \sum_k \mu_k^2 \langle\psi|m_k|\psi\rangle - \bar{\mu}^2 & (5)\\ &= \langle\psi|M^2 - \bar{\mu}^2|\psi\rangle + \langle\psi|\Delta M^2|\psi\rangle & (6)\end{aligned}$$

where we introduce the uncertainty operator

$$\Delta M^2 = \sum_k \mu_k^2 m_k - \bar{\mu}^2 \tag{7}$$

The first term in eq. (6) is the variance of μ_k which would arise if one was to "measure the observable M" in the usual sense. Note that there are always some states $|\psi\rangle$ for which this term vanishes. The second term is the additional uncertainty which arises because one is measuring a POVM and not a PVM. It does not depend on the average value $\bar{\mu}$. The uncertainty operator ΔM^2 thus characterizes the extra noise coming from the POVM. It has the following important properties:
1)*Positivity:* $\Delta M^2 \geq 0$ is a positive operator;
2) *Vanishing on PVM's:* $\Delta M^2 = 0$ if $\mathcal{M}$ is a PVM;
3)*Strict positivity on POVM's:* if $\mathcal{M}$ is not a PVM, then there exists a choice of μ_k such that $\Delta M^2 > 0$ is strictly positive (simply take $\mu_k = \delta_{kk_0}$ to be zero except for one value k_0, with k_0 such that m_{k_0} is not a projector);
4)*Reduction to classical random variables:* if the POVM elements are all proportional to the identity $m_k = |m_k|\mathbb{1}$, then the probabilities of the outcomes are independent of the quantum state, and μ_k is a classical random variable, with associated probabilities $p_k = |m_k|/d$ where d is the dimension of the Hilbert space. In this case the first term in eq. (6) vanishes and $\Delta M^2 = \mathbb{1}\sum_k(\mu_k - \bar{\mu})^2 p_k$ is the variance of the classical random variable μ_k.
5)*Additivity under tensor product:* Consider two POVM's and their associated values $\{m_i, \mu_i\}$ and $\{n_j, \nu_j\}$ acting on different systems. We construct the tensor product POVM as $\{m_i \otimes n_j, \mu_i + \nu_j\}$, where we associate to each outcome (i, j) the sum of the values $\mu_i + \nu_j$ for each outcome. Then the uncertainty operator for the tensor product POVM is $\Delta(M \otimes N)^2 = \Delta M^2 \otimes \mathbb{1} + \mathbb{1} \otimes \Delta N^2$.
6)*Increase under convex combination:* Consider two POVM's and their associated values $\{m_i, \mu_i\}$ and $\{n_j, \nu_j\}$ acting on the same system. We construct the convex combination of these two POVM's by realizing the first POVM with probability p, and the second POVM with probability q $(p + q = 1)$ to obtain a POVM with elements pm_i and qn_j to which are associated the values μ_i and ν_j respectively. The uncertainty operator for the convex combination POVM is $\Delta(pM + qN)^2 = p\Delta M^2 + q\Delta N^2 + pq(M - N)^2$.

Properties 5 and 6 are in direct analogy with the way the variances of independent classical random variables behave under addition and convex combination. The only non trivial property is the first one. We give a simple proof based on Naimark's theorem. Naimark's theorem states that any POVM can be viewed as a PVM on an extended Hilbert space. Let us focus on the case where the POVM elements are rank 1: $m_k = |m_k\rangle\langle m_k|$ (the general case follows by taking several of the μ_k to have the same value). Then there exists an extended Hilbert space $\tilde{H}$ which is the direct sum of the system Hilbert space H (on which $\mathcal{M}$ acts) and an ancillary Hilbert space H': $\tilde{H} = H \oplus H'$, and a basis of the extended Hilbert space $|\tilde{m}_k\rangle$, with $\langle\tilde{m}_l|\tilde{m}_k\rangle = \delta_{kl}$, such that its restriction to the system space gives the POVM $\mathcal{M}$:

$$|\tilde{m}_k\rangle = |m_k\rangle + |m'_k\rangle \tag{8}$$

with $|m_k\rangle \in H$ and $|m'_k\rangle \in H'$. We have the relation

$$\langle m'_l|m'_k\rangle = \delta_{kl} - \langle m_l|m_k\rangle \tag{9}$$

which allows us to rewrite ΔM^2 as

$$\Delta M^2 = \left(\sum_k \mu_k |m_k\rangle\langle m'_k|\right)\left(\sum_l \mu_l |m'_l\rangle\langle m_l|\right) \tag{10}$$

which is a manifestly positive operator.

It is interesting to note that the uncertainty relation eq. (6) can also be derived from the Robertson inequality. To this end note that the Naimark extension is not unique: $|\tilde{m}_k\rangle = |m_k\rangle + |m'_k\rangle$ and $|\bar{m}_k\rangle = |m_k\rangle + i|m'_k\rangle$ are two valid Naimark extensions of the same POVM. Hence we can define two operators $\tilde{M} = \sum_k \mu_k |\tilde{m}_k\rangle\langle\tilde{m}_k|$ and $\bar{M} = \sum_k \mu_k |\bar{m}_k\rangle_{ext}\langle\bar{m}_k|$. Applying the Robertson inequality eq. (1) we have $\Delta\tilde{M}\Delta\bar{M} \geq \frac{1}{2}|\langle[\tilde{M}, \bar{M}]\rangle|$.

But if the quantum state $|\psi\rangle$ has support only in the system space H, then $\Delta\tilde{M} = \Delta\bar{M} = Var(\mu)^{1/2}$ and $\frac{1}{2}[\tilde{M},\bar{M}] = i\Delta M^2$, yielding the second term in eq. (6). (On the right hand side of the Robertson inequality one can add a term containing the anti commutator of the two operators. By adding it one also recovers the first term in eq. (6)).

Examples. A well studied example (for more details see [13] and references therein) is the covariant joint measurement of p and x of the form $m_{px} = D_{px}\frac{m_0}{2\pi}D^\dagger_{px}$ where D_{px} is the displacement operator in phase space and m_0 is a normalized state. The variance of the position estimate in state ρ is the sum of the variance of the position operator x in state ρ and in state m_0. These two terms correspond exactly to the two terms in eq. (6). A similar decomposition holds for the variance of p.

Another example is the joint measurement of the σ_z and σ_x component of a spin 1/2 particle which is mathematically equivalent to a joint measurement of the path a particle takes in a Mach-Zehnder interferometer and of the output port by which it will exit[15, 18]. (It is thus conceptually the same but mathematically simpler than the Einstein version of the two slit experiment in which one tries to simultaneously determine the slit the particle passed through and to see the interference pattern, see [8, 9]). Such a measurement will be described by a POVM with 4 outcomes: m_{zx}, where the label $z = \pm 1$ ($x = \pm 1$) corresponds to inferring that a measurement of σ_z (σ_x) would preferentially have given the $z = +1$ or $z = -1$ ($x = +1$ or $x = -1$) outcome. One such POVM has elements

$$m_{zx} = \frac{\mathbb{1}}{4} + z\frac{\cos\theta}{4}\sigma_z + x\frac{\sin\theta}{4}\sigma_x \ . \qquad (11)$$

If we use this POVM to estimate the z (x) component of the spin, then the associated operators are $Z = \sum_{z,x=\pm 1} z m_{zx} = \cos\theta\sigma_z$ and $X = \sum_{z,x=\pm 1} x m_{zx} = \sin\theta\sigma_x$. Thus one is indeed simultaneously estimating the z and x components of the spin, but with reduced sensitivity with respect to measuring the observables σ_z and σ_x separately. The associated uncertainty operators are

$$\begin{aligned}\Delta Z^2 &= \sum_{zx} z^2 m_{zx} - (\sum_{zx} z m_{zx})^2 = (1-\cos^2\theta)\mathbb{1} \\ \Delta X^2 &= \sum_{zx} x^2 m_{zx} - (\sum_{zx} x m_{zx})^2 = (1-\sin^2\theta)\mathbb{1}\end{aligned} \qquad (12)$$

which we can group in the relation $\Delta X^2 + \Delta Z^2 = \mathbb{1}$ which implies the constraint on the sum of the variances

$$Var(Z) + Var(X) \geq 1 \ . \qquad (13)$$

It has been shown[18], and verified experimentally [23, 24], that eq. (13) holds for measurement schemes in which one first carries out a weak measurement of the path of the particle by correlating it with a probe and then measures the output port by which the particle exits. But such sequential measurements are only a small subset of all possible measurements. We now show that eq. (13) holds much more generally.

We consider POVM's that have 4 outcomes $z, x = \pm 1$, where outcome $z = +1, x = +1$ corresponds to guessing that measurements of σ_z and of σ_x would both have given outcome $+1$; and similarly for the other values of z, x. We require that the POVM be *unbiased*, by which we mean that, averaged over all input states, the probabilities of the different outcomes z, x are all equal. This implies that the POVM elements can be written as $m_{zx} = \mathbb{1}/4 + \vec{v}_{zx}\cdot\vec{\sigma}$ with $|\vec{v}_{zx}| \leq 1/4$ and $\sum_{zx}\vec{v}_{zx} = 0$. In addition for the measurement to be *faithfull* we require that $\vec{v}_{++}$ has positive z and x components. This corresponds to requiring that the state ψ that maximizes the probability of getting outcome $++$ would give with high probability $\sigma_z = +1$ and $\sigma_x = +1$ if one were to measure these operators. Similar conditions hold for the other values of z and x. As before we associate with outcome z, x the values $Z = z, X = x$. It then follows (in fact from the unbiasedness condition alone) that

$$\Delta Z^2 + \Delta X^2 = \mathbb{1}(2 - 2|\vec{v}_{++} - \vec{v}_{--}|^2 - 2|\vec{v}_{+-} - \vec{v}_{-+}|^2) \geq \mathbb{1}$$

which implies eq. (13). This result puts the uncertainty relation eq. (13) for *joint estimates* of which path/which output port a particle would take on the same level of generality as the uncertainty relations [16, 17] for *predictions* of which path/which output port the particle would take. Note that equality in eq. (13) is obtained if and only if $v_{++} = -\vec{v}_{--}$, $\vec{v}_{+-} = -\vec{v}_{-+}$ and $|\vec{v}_{zx}| = 1/4$. Thus the unbiased POVM with the least uncertainties can be realized by measuring with probability 1/2 the operator $\vec{v}_{++}\cdot\vec{\sigma}$ and with probability 1/2 the operator $\vec{v}_{+-}\cdot\vec{\sigma}$.

Entropic uncertainty relations. An alternative approach toward understanding the amount of randomness generated by a POVM is to lower bound entropy of the outcomes $H(\mathcal{M}) = -\sum_k p(k)\log_2 p(k)$ with $p(k) = \langle\psi|m_k|\psi\rangle$. The simplest such bound is given by the largest eigenvalue of the POVM elements m_k:

$$H(\mathcal{M}) \geq -\log_2\left(\max_{k\psi}\langle\psi|m_k|\psi\rangle\right) \ . \qquad (14)$$

Thus, except in the trivial case when one of the POVM elements is a projector, the entropy of the POVM outcomes is always positive. In some cases we have been able to improve this bound.

When the POVM can be realized by carrying out with probability 1/2 one of two non degenerate PVM's $\{|a_k\rangle\}$ and $\{|b_l\rangle\}$ (with $\langle a_k|a_{k'}\rangle = \delta_{kk'}$ and $\langle b_l|b_{l'}\rangle = \delta_{ll'}$) then eq. (2) implies

$$H(\mathcal{M}) \geq 1 - \frac{1}{2}\log_2\max_{k,l}|\langle a_k|b_l\rangle|^2 \ . \qquad (15)$$

Applied to the POVM eq. (11) this yields the bound

$$H(\{m_{zx}\} \geq 1 - \log_2\cos\theta \quad , \quad 0 \leq \theta \leq \pi/4 \qquad (16)$$

which is tight when $\theta = 0$ or $\theta = \pi/4$ but is suboptimal in between. Thus as the POVM $\{m_{zx}\}$ goes from estimating only σ_z ($\theta = 0$) to estimating σ_z and σ_x with equal sensitivity ($\theta = \pi/4$), the minimum entropy generated by the POVM increases from 1 to 3/2 bits. This is a particular example of what we expect is a general trade off: POVM's that probe more uniformly the Hilbert space (which can be a useful property for information processing) will generate more randomness (a deleterious property).

We have also improved on eq. (14) using a different method. Let consider bounds on the entropies of two POVM's, $\mathcal{M} = \{|m_k\rangle\langle m_k|\}$ and $\mathcal{N} = \{|n_l\rangle\langle n_l|\}$, whose elements are all rank 1, acting on the same state. As was noted in [25], the proof of the entropic uncertainty relation given in [6] immediately generalizes to:

$$H(\mathcal{M}) + H(\mathcal{N}) \geq -\log_2 \max_{kl} |\langle m_k | n_l \rangle| \ . \qquad (17)$$

This can be strengthened by noting that the Naimark extension is not unique: any PVM of the form $U'|\tilde{m}_k\rangle = |m_k\rangle + U'|m'_k\rangle$ with U' acting only on the ancillary Hilbert space H' is a possible extension of the POVM $\mathcal{M}$. Applying eq. (2) to the Naimark extension of the POVM's $\mathcal{M}$ and $\mathcal{N}$ and taking the best such bound yields:

$$H(M) + H(N) \geq \max_{U'} -\log_2 \max_{kl} |\langle \tilde{m}_k | U' | \tilde{n}_l \rangle| \ . \qquad (18)$$

Obviously eq. (17) is a particular case of eq. (18).

We can now go back to the case of a single POVM. Let us take in eq. (18) the two POVM's to be identical but with different Naimark extensions. This yields the entropic uncertainty relation for a single POVM:

$$H(\mathcal{M}) \geq \max_{U'} -\frac{1}{2} \log_2 \max_{kl} |\langle \tilde{m}_k | U' | \tilde{m}_l \rangle| \ . \qquad (19)$$

As an illustration we consider the POVM described in [21] which acts on the symmetric space of two spin 1/2 particles and is composed of 4 elements, each proportional to the projector onto two parallel spins oriented along the 4 corners of a tetrahedron. The Naimark extension of this POVM can be constructed explicitly as $|\tilde{m}_0\rangle = \frac{\sqrt{3}}{2}|\uparrow_z\rangle|\uparrow_z\rangle + \frac{1}{2}|a\rangle$, $|\tilde{m}_j\rangle = \frac{\sqrt{3}}{2}|\uparrow_j\rangle|\uparrow_j\rangle - \frac{1}{2}|a\rangle$ where $|\uparrow_j\rangle = (|\uparrow_z\rangle + \sqrt{2}e^{i2\pi j/3}|\downarrow_z\rangle)/\sqrt{3}$, $j = 1, 2, 3$. If we take $U'|a\rangle = -|a\rangle$ (with U' acting as the identity on the space of the two spins), then eq. (19) implies that the entropy of this POVM is bounded by $H(\mathcal{M}) \geq 1$ bit, which is significantly better than the bound $H(\mathcal{M}) \geq -\log_2(3/4)$ which follows from eq. (14).

Conclusion In this work we have studied of the amount of randomness generated by a POVM. Our first approach gives a bound on the variance of a variable associated to the POVM outcomes. We have used this method to generalize the domain of applicability of the inequality proposed by B.-G. Englert for the uncertainty relation governing joint measurements of which path / interference visibility. The second method consists in bounding the entropy of the POVM outcomes, and has the same attractive properties as the entropic uncertainty relation for PVM's. We give examples of such bounds. All the uncertainty relations for POVM's we present can in fact be derived from the uncertainty relations (1) and (2) for PVM's. This is quite natural since we are always exploring reformulations of the same physical property: in quantum mechanics non commuting quantities can never simultaneously take definite values.

We acknowledge support by EU project QAP contract number 015848, and by the IAP project P6/10 Photonics@be.

[1] W. Heisenberg, Z. Phys. **43** (1927) 172; English translation in J. A. Wheeler and W. Zurek, *Quantum Theory and Measurement*, Princeton University Press, 1983
[2] H. P. Robertson, Phys. Rev. **34** (1929) 163
[3] E.H. Kennard, Z. Phys. **44** (1927) 326352.
[4] D. Deutsch, Phys. Rev. Lett. **50** (1983) 631
[5] K. Kraus, Phys. Rev. D **35** (1987) 3070
[6] H. Maassen and J. B. M. Uffink, Phys. Rev. Lett. **60** (1988) 1103
[7] I. Bialiniki-Birula and J. Mycielsky, Commun. Math. Phys. **44** (1975) 129
[8] W. H. Wooters and W. K. Zurek, Phys. Rev. D **19** 473 (1979)
[9] L. S. Bartell, Phys. Rev. D **21** 1698 (1980)
[10] B.-G. Englert and K. Wódkiewicz, Phys. Rev. A **51**, R2661(1995).
[11] M. Ozawa, Ann. Phys. **311**, 350 (2004)
[12] R. F. Werner, in *Quantum Information Statistics and Probability*, edited by O. Hirota, Paramus, N. J. Rinton, pp. 153-171; quant-ph/0405184
[13] P. Busch, T. Heinonen and P. J. Lahti, quant-ph/0609185
[14] J. Kiukas, P. Lahti and K. Ylinen, J. Math. Phys. 47, 072104 (2006).
[15] P. Busch and C. R. Shilladay, Phys. Rev. A **68** (2003) 034102
[16] D. M. Greenberger and A. Yasin, Phys. Lett. A **128**, 391 (1988).
[17] G. Jaeger, A. Shimony, and L. Vaidman, Phys. Rev. A **51** (1995) 54
[18] B.-G. Englert, Phys. Rev. Lett. **77** (1996) 2154
[19] S. Dürr, Phys. Rev. A **64** (2001) 042113
[20] A. Peres and W. K. Wootters, Phys. Rev. Lett. **66** (1991) 1119
[21] S. Massar and S. Popescu, Phys. Rev. Lett. **74** (1995) 1259
[22] S. Massar and S. Popescu, Phys. Rev. A 61, 062303 (2000); A. Winter and S. Massar, Phys. Rev. A 64, 012311 (2001); A. Winter, Comm. Math. Phys. 244, 157 (2004).
[23] S. Dürr, T. Nonn, and G. Rempe, Phys. Rev. Lett. **81**

(1998) 5705
[24] P. D. D. Schwindt, P. G. Kwiat, and B.-G. Englert, Phys. Rev. A **60** (1999) 4285
[25] Michael J. W. Hall, Phys. Rev. A **55** 100 (1997)